

SUBSIDE

Subsidence System for Insight and Data Exploration

A Resource for Texas Groundwater Management Areas

*A collaboration with the Texas Water Development Board
with*

*Texas Advanced Computing Center, U.S. Geological Survey, The University of Texas at Arlington, with contributions from the
subsidence community in the State of Texas*

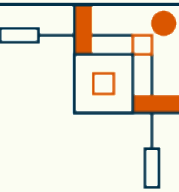
What is SUBSIDE?

A new resource to help Texans access and analyze subsidence information

Built on Information shared by the subsidence and groundwater management community in Texas, SUBSIDE brings together multiple data sources in one place, such as

- InSAR satellite imagery from NASA's OPERA program
- GPS/GNSS ground station measurements
- Well data and geological information
- Offers ready-to-use analytical tools through shareable computational cookbooks
- Provides a searchable data catalog for discovering available datasets
- Designed to support local groundwater planning and management

Why This System?



The Opportunity

- Subsidence data exists across multiple agencies
- Stakeholders have requested easier access to information
- Advanced analysis tools can now be made more accessible
- Technology can help extend technical capacity statewide

The Value

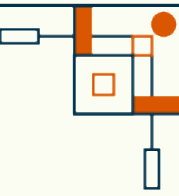
- Better information supports proactive infrastructure planning
- Coastal and inland communities benefit from integrated data
- Groundwater managers gain additional tools for decision-making
- Supports the state's commitment to sound water planning

The Approach

- Partner with TACC for secure, scalable infrastructure
- Leverage existing federal data investments (USGS, NASA)
- Build reusable tools that stakeholders can adopt and access
- Provide training and support for local implementation

Initial Design Duration: October 2024 – June 2026

Progress to Date



~85%

Project Milestones

6 Computational Cookbooks

Ready for stakeholder use

Data Portal Online

Browse and download datasets

InSAR Processing

Statewide satellite coverage

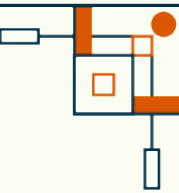
Metadata Standards

Consistent data documentation

What's Now Available

- Secure storage and computing resources allocated on TACC systems
- Comprehensive review of existing subsidence monitoring programs
- Best practices framework for data governance (90% complete)
- Metadata guidelines developed in coordination with TWDB and USGS

How This May Support GMA-12



Understanding Your Aquifers

TWDB subsidence vulnerability assessment for GMA-12:

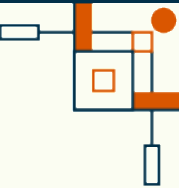
Aquifer	Vulnerability Index
Yegua-Jackson	5.9
Brazos River Alluvium	5.3
Carrizo-Wilcox	4.7
Queen City	4.2
Sparta	4.2

These indices help identify where monitoring data may be most valuable for local planning.

Resources Available to You

- Satellite-based subsidence data
 - Rates in mm/year for your region
 - Historical trends going back 7+ years
- Analytical tools you can use directly
 - No specialized software required
 - Step-by-step workflows
- Connection to existing planning tools
 - Can complement GAM modeling efforts
- Data discovery portal
 - Find what's available for your area

Looking Ahead: Expanding Capabilities



Our goal is to ensure that every groundwater management area has access to the same quality of information and analytical tools.

Your input helps shape SUBSIDE

Monitoring Support

- Potential to flag unusual patterns for local review
- Could provide early indicators for infrastructure planning
- Designed to complement existing local monitoring

Scenario Exploration

- Easier access to model scenario comparisons
- Tools to explore what-if questions locally
- May help inform DFC discussions

Data Integration

- Connecting subsidence trends with water level data
- Potential links to regional planning documents
- Building toward more integrated views

We welcome your feedback on which capabilities would be most valuable for your work.

Analyze Your Own Aquifer Data

Bring your own data and questions. Leave with the tools to answer them.

Training Institute • May 11-15, 2026 • TACC, Austin

After this workshop, you will be able to:

- Run subsidence analyses for your specific region using SUBSIDE tools
- Access and interpret InSAR satellite data for your aquifers
- Use computational cookbooks to create reproducible, shareable workflows
- Connect your local monitoring data with statewide resources

To Learn More and Register

Contact Jean Perez
Jean.Perez@twdb.texas.gov

Supplemental

```
[*]: sim.write_simulation()
success, buff = sim.run_simulation()
if not success:
    print("Something bad happened.")
```

Government shall be held liable for any damages resulting from the authorized or unauthorized use. Also refer to the USGS Water Resources Software User Rights Notice for complete use, copyright, and distribution information.

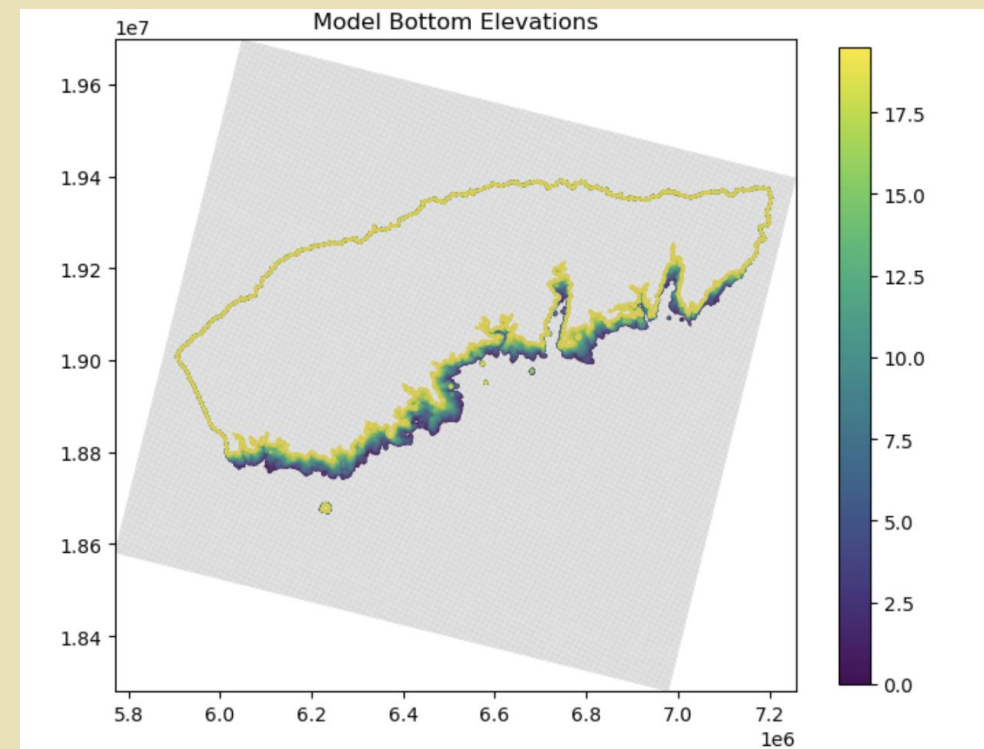
MODFLOW runs in SEQUENTIAL mode

Run start date and time (yyyy/mm/dd hh:mm:ss): 2026/04/01 15:57:05

Writing simulation list file: mfsim.lst

Using Simulation name file: mfsim.nam

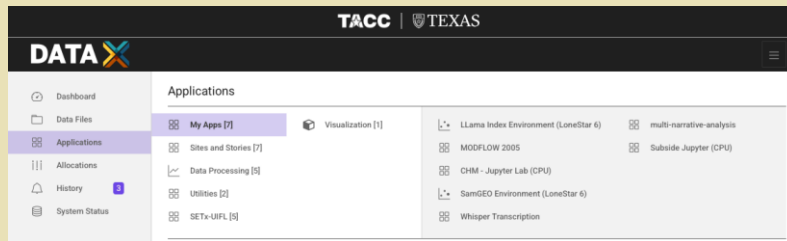
Solving:	Stress period:	1	Time step:	1
Solving:	Stress period:	2	Time step:	1
Solving:	Stress period:	3	Time step:	1
Solving:	Stress period:	4	Time step:	1
Solving:	Stress period:	5	Time step:	1
Solving:	Stress period:	6	Time step:	1



Dataset Workflow

This proof-of-concept workflow highlights a basic analysis

Selecting App in portal >>



Analyze data & create products &

```
[1]: # Import required libraries
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
from sklearn.feature_extraction.text import TfidfVectorizer
from nltk.corpus import stopwords
import matplotlib.pyplot as plt
import seaborn as sns

# Load your cleaned data
df = pd.read_pickle('/Users/samp33/BittenMLP/output/cleaned_globalSubDecember24Bibliography.pkl')

print("Data loaded successfully!")
print("Number of entries: {}".format(len(df)))

Data loaded successfully!
Number of entries: 8489

[2]: # Create combined text field
df['combined_text'] = df['abstract'].fillna('') + ' ' + df['keywords'].fillna('') + ' ' + df['title'].fillna('')

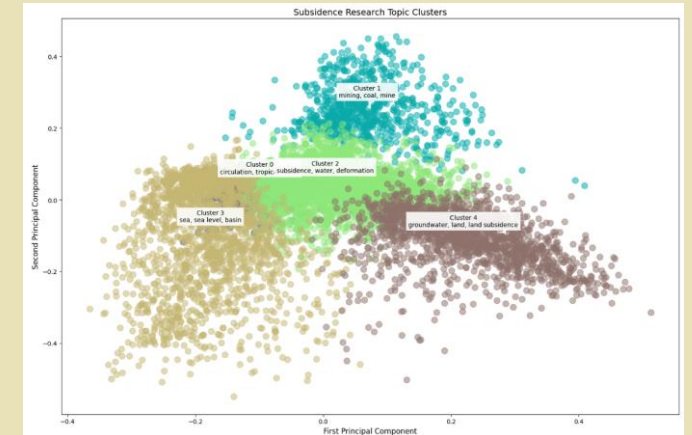
# Create TF-IDF matrix
vectorizer = TfidfVectorizer(
    stop_words=stopwords.words('eng'),
    max_features=1000,
    ngram_range=(1, 2)
)

# Process the text data
tfidf_matrix = vectorizer.fit_transform(df['combined_text'])
feature_names = vectorizer.get_feature_names_out()

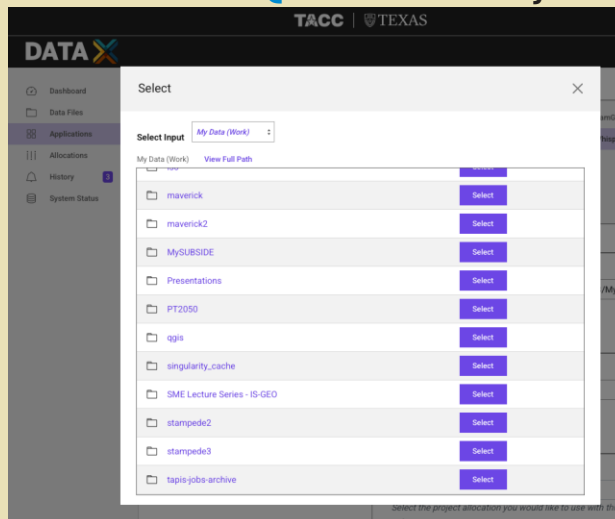
print("TF-IDF matrix created successfully!")
print("Number of features: {}".format(tfidf_matrix.shape[1]))

TF-IDF matrix created successfully!
Number of features: 1000
```

Generate Interactive Graphics >>



Submit to Queue to run job or analysis >>



Parameters

min_speakers (appArgs) **Required**

1

Minimum number of speakers.

Configuration

Allocation **Required**

PT2050-DataX

Select the project allocation you would like to use with this job submission.

Queue **Required**

gpu-a100-dev

Select the queue this job will execute on.

Maximum Job Runtime (minutes) **Required**

100

The maximum number of minutes you expect this job to run for. Maximum possible is 120 minutes. After this amount of time your job will end. Shorter run times result in shorter queue wait times.

Output

- Step 1**
 - Total records (n = 19,000)
 - Databases searched: Web of Science & Semantic Scholar
 - Language: English Only, Timeframe: Through January 31, 2025
- Step 2**
 - Unique articles (n = 8,009)
 - Excluded: duplicates, unrelated, Other literature
 - Created bibtext data collection, Conducted full text review with bibliometric tools
- Step 3**
 - Topical Analysis on Peer Reviewed Data collection
 - Dictionary, cluster analyses, and theme identification
- Step 4**
 - Multi-stage filtering and AI-supported queries
 - Characterizing Subsidence Processes and Datasets of Interest
 - Query: Information needs for change tracking
- Step 5**
 - Multi-stage filtering and AI-supported queries
 - Evaluating analytical workflows, simulation modeling, and software tools
 - Queries: Comparative analysis and integrating AI tools with subsidence information.
- Step 6**
 - Preparing Summaries and Figures
 - Aligning the results with requirements, capabilities, and feature design specifications

Share via portal or website

